(12)

NOSC TR 1123

NOSC

NAVAL OCEAN SYSTEMS CENTER San Diego, California 92152-5000

Technical Report 1123
July 1986

Broadband Sonar Classification Cues

An Investigation

D. W. Martin
W. W. L. Au

Approved for public release; distribution is unlimited.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): NOSC TR 1123	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION: Naval Ocean Systems Center; 6b OFFICE SYMBOL (if applicable): Codes 513 and 512	7a NAME OF MONITORING ORGANIZATION
6c ADDRESS (City, State and ZIP Code): San Diego, CA 92152-5000	7b ADDRESS (City, State and ZIP Code)
8a NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b OFFICE SYMBOL (if applicable): ONR-251	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State and ZIP Code): Alexandria, VA 22217	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 62759N; PROJECT NO F59558; TASK NO MM33; AGENCY ACCESSION NO DN 488838

11 TITLE (include Security Classification)
Broadband Sonar Classification Cues, An Investigation

12 PERSONAL AUTHOR(S)
D. W. Martin and W. W. L. Au

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM ____ TO ____	July 1986	44

16 SUPPLEMENTARY NOTATION

17 COSATI CODES

FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Echo Discrimination

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The human auditory system has excellent pattern recognition capabilities and can be used to identify acoustic cues in broadband sonar classification tasks. Sonar echo-discrimination experiments were conducted with human subjects to (1) measure human auditory discrimination performance using broadband sonar echoes from geometric targets, (2) determine the acoustic cues used by subjects in discriminating between targets, (3) develop software algorithms to extract echo features similar to those used by humans, and (4) determine whether the features can be used for automatic target classification.

Human subjects discriminated between a variety of target echoes, with scores exceeding 90-percent correct on most tasks, using mostly time-domain cues including time-separation pitch, rise time, and duration differences. Time-domain feature extraction software was developed, with echo features statistically representing the time-separation and amplitude of echo highlights. An automatic pattern recognition algorithm using these features obtained 90-percent correct performance on a seven-alternative material composition discrimination task with chance performance being 14-percent correct.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
D.W. Martin	(619) 254-4443	Code 513

DD FORM 1473, 84 JAN

83 APR EDITION MAY BE USED UNTIL EXHAUSTED
ALL OTHER EDITIONS ARE OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

CONTENTS

ILLUSTRATIONS

TABLES

INTRODUCTION

Broadband sonar echoes convey target information not available in echoes from narrowband sonars. Traditional signal processing methods provide poor target recognition information for low Doppler targets. Mathematical descriptions of echoes based on scattering theory (references 1 and 2) are often specific to individual targets and cannot account for slight changes in target parameters. Target recognition methods such as acoustic imaging and T-matrix formulations require processing of unwieldy amounts of information. A slight change in target characteristics can require yet more information for reprocessing. Much of the information is redundant or would only be important in a context-independent recognition task. A smaller set of discrimination cues is needed, with the system retaining only that information necessary to discriminate among a limited set of targets within a particular context. We do not have to form complete images of targets, but only to have cues that uniquely characterize targets likely to occur. For example, detecting an edge is sufficient to discriminate between a sphere and a cube; no other information is required if this is the only discrimination of interest. If a priori knowledge exists about the context of a sonar task, a subset of back-scattered information should be sufficient for target identification problems. A set of cues is needed that could identify target shape, independent of material, size, or aspect, or target material independent of other parameters.

The human auditory system has excellent pattern recognition capabilities and can identify acoustic cues useful in broadband sonar classification tasks. Humans do not have perfect memories for signals, but they can be trained to adaptively attend to a small set of relevant cues. Computer algorithms, on the other hand, cannot even approximate human performance in speech recognition, including voice-independent word recognition and word-independent voice recognition. Human performance with nonspeech signals is also excellent (references 3, 4, and 5).

This study's objectives and a background discussion on human auditory pattern recognition are in the next two sections. The background discussion is presented to familiarize the reader with concepts of auditory pattern recognition relating to sonar discrimination. These concepts, which provide the foundation for our experiments, have not been systematically presented previously. The experimental methods used to measure echo discrimination performance and a description of the parameters of some target echoes that are potentially relevant to the recognition problem are discussed in subsequent sections. Results from four echo discrimination experiments are presented and discussed next. The section on feature extraction and pattern recognition describes a method of extracting from target echoes acoustic features (similar to features used by humans) that were used in a pattern recognition algorithm to classify targets. Classification performance with these features is compared to that of an earlier feature extraction/pattern recognition algorithm developed by Chestnut et al. (reference 6). Finally, other possibilities for using concepts from human pattern recognition to guide signal processing efforts are outlined.

OBJECTIVES

1. Measure human auditory discrimination performance using broadband sonar target echoes.

2. Identify the acoustic cues used by subjects to discriminate target shape, material composition, and internal structure, as well as identify useful cues for aspect-independent target discrimination.

3. Develop software algorithms to extract echo features similar to those used by humans.

4. Determine whether echo features can be used for classification using automatic pattern recognition algorithms.

BACKGROUND

The methods by which humans acquire information, extract features, and determine which features are important in a pattern-recognition decision are not well understood. We assume that subjects somehow encode a perceived stimulus as a set of features and structural relations among features (reference 7). This set of features is then compared to stored patterns or templates in memory, and a matching pattern is chosen from the subset of memorized patterns based on perceived stimulus similarity (reference 7). The subject may have only partial information about the perceived stimulus, or the memory image may be incomplete. Many recognition models assume that features in memory are forgotten independently and that the perceived relevance of a feature can affect the decay rate (reference 7). Eventually, only the most important features remain to describe a pattern. Perceived structural relations among features and a subject's previous listening experience can affect the detectability of each component of a feature list (references 8 and 9). However, to discriminate between two patterns, a subject must have an opportunity to detect at least one feature describing the difference between patterns (reference 10).

Investigators have used similarity judgments or confusion matrices to identify features for auditory discrimination of complex sounds (references 4, 5, and 11). Similarity judgments and confusion matrices for the same stimuli identify the same stimulus dimensions as important. If the number of dimensions along which two stimuli differ is increased, they will be judged less similar and will be confused less often. Subjects differ in their judgments of which cues are most important on a given discrimination task. Howard (reference 12) found that the degree to which a given feature contributed to a similarity judgment was strongly influenced by the categories into which the experimenter partitioned the stimulus set. The subjects' ability to group a set of dissimilar stimuli into a particular class requires the emphasis of some features and the de-emphasis of others. Training in this area is critical in sonar discrimination tasks if target echoes are to be grouped into generic target classes. In many sonar tasks, naive subjects should be able to discriminate between targets, but generalizing this discrimination to "target recognition" will require extensive training. Identifying a simple set of cues should enhance this process.

IN-AIR SONAR

Many different animals, including humans (reference 13), bats (reference 14), and some species of birds (reference 15) use in-air sonar. Blind people use self-generated signals to detect and avoid obstacles (reference 13). Both broadband clicks and hisses are superior to narrowband tonal signals for obstacle avoidance (references 13 and 16). Learning is sudden and insightful, implying that subjects need to recognize the existence of a previously unused perception (reference 17). The obstacle perception

seems to depend on a rise in perceived echo pitch as obstacles are approached (references 16 and 18). Loudness changes are insufficient for obstacle detection. although they may be involved in size or material-composition discriminations based on target strength (references 18 and 19). Some subjects can make simple material or shape discriminations, e.g., wood versus metal or square versus circle, and can discriminate different sized objects (references 13 and 19). However, the performance of the human sonar system is inferior to those of many animals and electronic devices.

Continuous-transmission frequency modulation (FM) sonars with auditory displays for in-air target detection and discrimination by the blind have been designed by Kay (reference 20). Echoes from a stationary flat surface are displayed as pure tones whose frequency decreases with decreasing target range. If the target has shape or texture features that are large compared to a wavelength, a complex tonal structure is heard (references 20 and 21). The complex stimuli received from various classes of objects can be remembered and are often generalized to include new objects of a class.

UNDERWATER-SONAR

The information that can be extracted from target echoes in water is different from that in air. In water, the acoustic signals can penetrate into targets so that aural discriminations of material composition and internal object structure become possible (reference 3). Sonar discrimination experiments have been performed with dolphins, humans, and electronic systems. In this section, only dolphin and human studies will be discussed.

In experiments concerning target size, dolphins discriminated, with 100-percent correct performance, between solid steel spheres 5.4 and 6.35 cm in diameter (reference 22) and between hollow aluminum cylinders with diameters of 7.6 and 6.35-cm (reference 23). Differences in time-separation pitch related to highlight spacing in echoes from the cylinders were probably the most salient cue. Differences in echo intensity may also have contributed to size discriminations.

In an experiment concerning target shape, a dolphin discriminated between cylinders and cubes independent of target aspect, except when the flat top of the cylinder was facing the animal (reference 24). These discriminations may be based on angular variations in target strength and on the dolphin's perception of multiple echoes from target edges. A dolphin discriminated between foam spheres and cylinders with performance exceeding 90-percent correct (reference 25).

Material composition discrimination between dimensionally identical cylinders was performed by a dolphin in the following cases: aluminum versus rock, aluminum versus steel, and aluminum versus bronze (reference 23). The dolphin's discrimination of aluminum and glass cylinders was 80-percent correct for 3.8-cm-diameter targets and chance for the 7.6-cm-diameter targets (reference 26). Glass and aluminum have nearly identical acoustic impedances and sound velocities Hammer and Au (reference 23) also found that dolphins could discriminate between dimensionally identical aluminum cylinders differing in internal structure. The dolphin also discriminated between hollow and solid aluminum cylinders, and between water-filled cylinders differing only in wall thickness (reference 23).

The dolphin's 50-percent detection threshold for a 7.6-cm-diameter water-filled sphere occurred in ambient noise at a range of 113 meters (reference 27). Au and Turl (reference 28) obtained good detection performance with water-filled aluminum cylinders in front of a clutter screen.

Results from human listening experiments indicate that the human auditory system is an excellent pattern recognizer and that broadband sonar echoes can supply information sufficient for discrimination of many types of targets. Fish et al. (reference 3) trained divers to discriminate between various plates 1 meter in front of them, using a head-coupled sonar that emitted broadband ultrasonic pulses and digitally stretched the echoes. Subjects discriminated between plates varying in shape (squares, circles, and triangles), material (copper, brass, and aluminum), and thickness. The divers' performance was between 80-percent and 100-percent correct. Diercks et al. (reference 29) used broadband FM echoes at five bandwidths to measure human discrimination performance between solid and hollow metal spheres and cylinders. Subjects reported that the rate of amplitude fluctuation during the echoes was a useful cue for discriminating target wall thickness, but only for signals having the widest bandwidth. Sphere-cylinder discriminations were based on a slower rise time for the sphere echoes and an amplitude notch shortly after sphere echo onset.

EXPERIMENTAL METHODS

This section describes procedures used to measure human discrimination of target echoes and to identify salient discrimination cues. Tasks included material composition discrimination, sphere-cylinde discrimination, aspect-independent cylinder discrimination, and target detection and discrimination in noise.

PROCEDURE

Echoes were obtained using broadband ultrasonic pulses similar to dolphin echolocation pulses. The incident signal shown in figure 1 had a 120-kHz center frequency and a 3-dB bandwidth of approximately 39 kHz. Targets were suspended in a saltwater pool 2.4 meters from the transducer. Echoes were digitized at a 1-MHz sample rate and recorded on magnetic tape. The taped data were later transferred to a PDP-11/40 computer. The echoes were played to human subjects through headphones in a sound booth. Four pulses were played per second at one fiftieth of the original sample rate. Thus, the stimuli that the subjects heard had peak frequencies of about 2.4 kHz and durations 50 times greater than the original echoes. The PDP-11 computer controlled the selection and playback of echoes, recorded subjects' responses, generated correct-response feedback after each trial, and displayed a summary output after each session. The experimental test conditions are shown in figure 2.

On each trial of a 64-trial session, subjects indicated whether the echoes they heard resulted from target class A or B by pressing one of two buttons on a response box, marked A and B, respectively. After each response, a light indicating the correct answer was illuminated for 2 seconds. Before each session, subjects practiced the echo discriminations by pressing the buttons to generate sample echoes from each of the target classes.

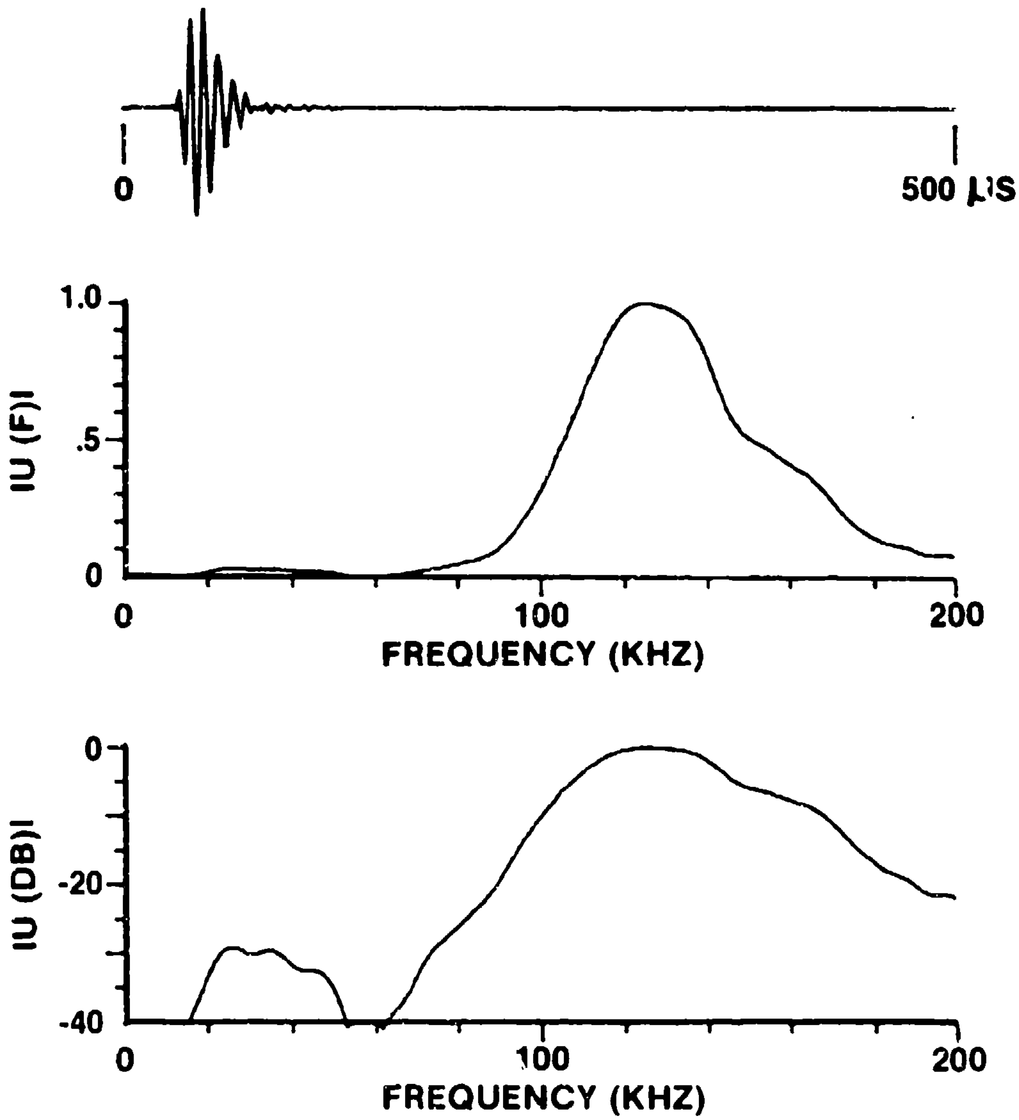

Figure 1. Simulated dolphin echolocation signal used as the incident signal. The top trace is the time-domain representation of the incident signal. The bottom traces are the frequency domain representations of the signal in linear and logarithmic scales, respectively.

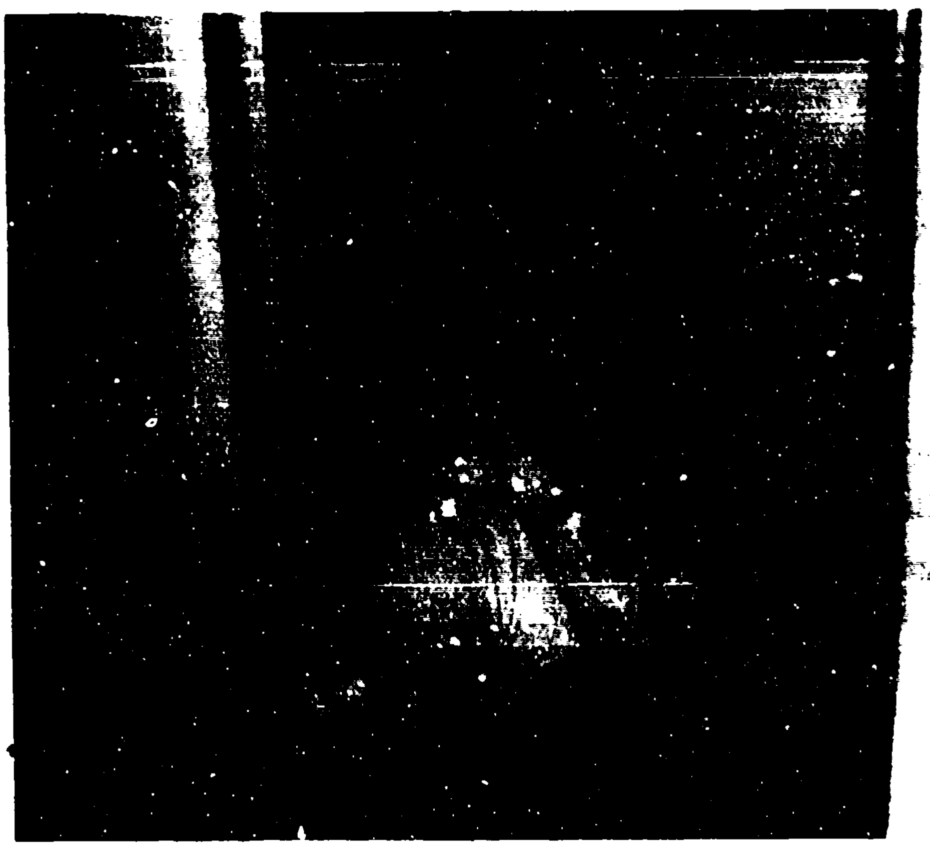

Figure 2. Experimental configuration for listening test: (top) PDP-11/40 computer system, and (bottom) subject in sound isolation booth.

A target class could include one or more targets. For example, in a sphere-cylinder discrimination, class A could contain several different spheres. Subjects were instructed to discriminate between target classes, even when a class contained multiple targets. Each target was represented by 10 echoes; there were often small variations between echoes for a given target. On each trial, a single echo was repeated until the subject responded A or B. Different echoes from the same target could occur on later trials. A single echo was repeated on each trial to prevent subjects from making A-B classifications based on the echo-echo variability for a given target. For each trial, individual echoes from the 10 were chosen at random. Target presentation was also randomized with the constraint that echoes from the same target could not occur on more than three consecutive trials, and all targets occurred an equal number of times within a session.

After each session, subjects described the cues used to make discrimination decisions. In the modeling phase of this study, these cue descriptions were used to guide the modification of echoes, producing signals with either enhanced or degraded cues. Cue descriptions were used rather than similarity judgments or confusion matrices, because the dimensions that described the differences between target echoes were not known in advance. In general, cues useful for target recognition could not be found by inspection or by known processing methods.

In all the discrimination tasks, variations in target strength between echoes were removed as cues. This was done because discriminations based on target strength differences could result from differences in target range instead of differences in target properties. Target strength cues were eliminated by normalizing the peak amplitudes of the echoes. Martin and Au (reference 30) found this procedure to be superior to normalizing total echo energy, because the subjects did not integrate over the entire duration of the signals. If they did, small temporal variations within an echo would be lost.

The methods described were common to all discriminations. Procedures relevant to particular tasks are described in the appropriate following sections.

DESCRIPTION OF ECHOES

Sample echo waveforms for solid and hollow aluminum cylinders are shown in figure 3. The waveforms contain highlights from multiple internal reflections, with the differences in highlight arrival times caused by different acoustic path lengths in the cylinders. Differences in the speed of sound in two materials or differences in target size will affect the arrival times of multiple-reflected components (reference 31). Multiple reflections along a given acoustic path through the target will be periodic with successively decreasing amplitudes. When targets are of similar size and composition, simple inspection of waveforms or spectra does not lead to accurate discrimination.

Human auditory discrimination of the echoes shown in figure 3 is based on differences in echo duration and in perceived time-separation pitch (reference 30). When two highly-correlated broadband pulses are separated by time (T), a time separation pitch (TSP) can be perceived with a frequency 1/T, whether or not the signal contains significant spectral energy at this frequency (references 32, 33, and 34).

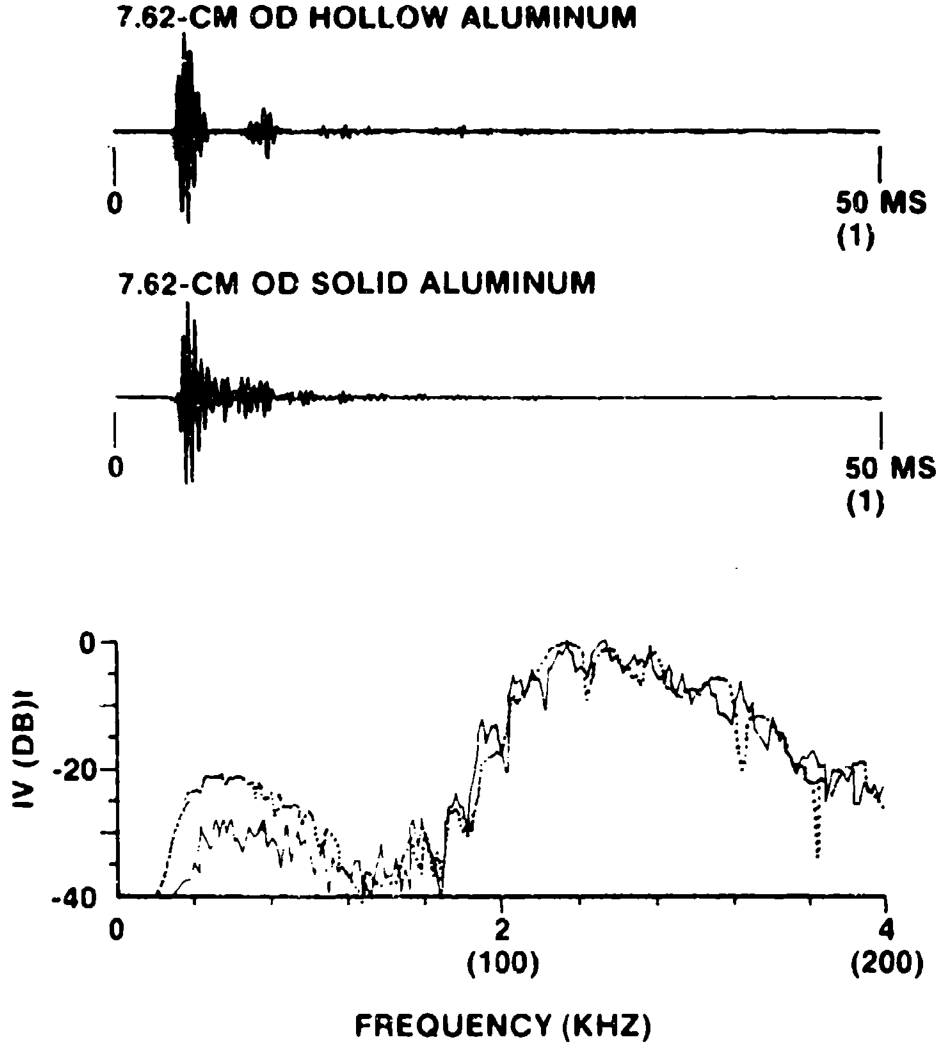

Figure 3. Typical echo waveforms and frequency spectra for the 7.62-cm-diameter hollow and solid aluminum cylinders. The solid line spectrum is for the hollow aluminum, and the dotted spectrum is for the solid aluminum.

TSP is believed to be a salient cue in the discrimination experiments involving differences in size or material composition. The time-separation pitch associated with target echoes cannot be easily verified using tone-matching experiments. Three or more echo components with variable separations and amplitudes produce a complex TSP, not necessarily defined as 1/T for any two components. The complex pitches associated with triple-pulse stimuli are discussed in a study by Ceruti et al. (reference 35), and the complexities introduced by differences in pulse amplitudes are discussed by Gillespie (reference 36).

In previous experiments with broadband sonar echoes (reference 30), subjects perceived echoes as clicks. Some echoes also included a leading or trailing hiss. Some clicks, such as solid aluminum echoes, had a metallic ringing sound, probably the result of periodicity within the echoes. The hissing sounds result from low-amplitude and uncorrelated echo components. Subjects also perceived spectral and rise time differences in the echoes.

Discrimination cues can often be predicted from geometrical acoustics For example, a duration cue might be used to discriminate between water-filled and air-filled aluminum cylinders. Echoes from the water-filled cylinder contain highlights from multiple internal reflections, whereas the metal-air interface appears as an infinite-impedance barrier, resulting in only a single dominant highlight for this echo. Cue descriptions will be included with performance results for each experiment.

EXPERIMENT I: MATERIAL COMPOSITION DISCRIMINATION

The material composition discrimination results and relevant discrimination cues were reported in detail by Martin and Au (reference 30) and are summarized here for completeness. Subjects discriminated between water-filled and solid 7.6-cm-diameter aluminum cylinders, with 98-percent correct responses, using both TSP and echo duration as cues. Performance on material composition discriminations using water-filled cylinders of aluminum, bronze, and steel of 3.8-and 7.6-cm-diameters exceeded 95-percent correct, with time-separation pitch differences as the primary cue. Discrimination between aluminum and glass cylinders of the same dimension showed large differences between subjects. Scores varied between 75-percent and 95-percent correct. Subjects with the best scores reported a longer duration for the aluminum echoes. Echoes seemed to damp out more quickly in the glass cylinders than in the aluminum.

A modified material discrimination experiment was performed in this study using the same targets as the previous study (reference 30). In the present study, two subjects discriminated between echoes from the same targets as above, after the echoes were passed through a replica correlator filter. For these subjects, all discriminations except the 7.6-cm-aluminum-glass cylinders resulted in the same performance levels as with unfiltered echoes. The aluminum-glass discrimination for these subjects averaged 92-percent correct for unfiltered echoes and 80-percent correct for filtered echoes. Subjects reported that when TSP contributed to the discrimination, the same cues were present in the unfiltered and filtered signals. Thus, correlation processing may have applications in high-noise environments to improve discrimination performance by removing uncorrelated noise from echoes. Hammer and Au (reference 23) examined the graphical outputs of this type of filtering process to identify possible cues used by dolphins. They found that the filter

responses for targets that were easily discriminable by the dolphin could also be easily discriminated from graphic displays. However, for targets that were difficult for the dolphin to discriminate, the filter responses were very similar.

The subjects' responses in the material discrimination tasks indicated that learning was sudden and insightful rather than gradual. For example, performance might remain at 65-percent correct for three sessions with the subject reporting confusion, and then improve to 90-percent correct for session four, remaining high in later sessions.

EXPERIMENT II: SPHERE-CYLINDER DISCRIMINATION

Discrimination was measured between spheres and cylinders of several different sizes made of foam, solid aluminum, and water-filled steel. All cylinder echoes were collected at broadside aspect. Tests were conducted using both two-target (one sphere and one cylinder) and four-target (two of each) conditions. Discrimination experiments were also conducted with foam target echoes modified by applying a time window to the signals. This time window eliminated an air-water interface reflected component from the echoes.

Foam targets and presentation schedules are in table 1. The same targets and combinations were used in similar experiments with dolphins (reference 25). Target sizes were chosen such that the target strengths of the two classes overlapped, eliminating target strength as a useful discrimination cue. The metal targets are described in table 2.

Discrimination results pooled across subjects for the foam targets are in table 3. The average of correct discrimination varied between 84- and 96-percent depending on the targets used. With one exception, variations in individual's scores were within 3 percent of their mean scores. For the comparison S1 and S2 versus C4 and C5, individual scores varied between 76- and 91-percent correct.

Subjects reported using two cues for these discriminations: a higher pitch for cylinder echoes and low-frequency reverberation in the sphere echoes. The pitch difference probably occurs because the target strength of a cylinder increases with frequency and is constant for a sphere. Because the foam targets do not have internal reflections, the observed pitch differences could not have resulted from TSP. The low-frequency reverberation in the sphere echoes probably resulted from reflection at the air-water interface. Au et al. (reference 25) attributed a dolphin's discrimination performance to the surface-reflected component. For tests with echoes that had no surface-reflected component, the subjects' discrimination performance dropped an average of 8 percent (table 3, windowed total). However, performance exceeded 80-percent correct on all tasks. The reverberation present in sphere echoes was helpful, but not necessary, for discrimination.

The results for two subjects discriminating metal spheres and cylinders are in table 4. Performance for all comparisons exceeded 94-percent correct. This experiment was conducted after the tests with foam targets. Subjects reported that the metal target echoes did not sound like the foam target echoes; internal reflections caused longer echo durations for the metal targets. However, subjects reported using the same discrimination cues: a higher pitch for cylinders and more reverberation for spheres.

Table 1. Foam targets and presentation schedules. The dimensions of the foam spheres (diameter) spheres and cylinders (diameter times length) are as used in the shape discrimination test.

Foam Targets

Spheres		Cylinders	
S1	10.2 cm	C1	1.9 x 4.9 cm
S2	12.7 cm	C2	2.5 x 3.8 cm
S3	15.2 cm	C3	2.5 x 5.1 cm
		C4	3.8 x 5.4 cm

Presentation Schedule

S2	C4
2 & S3	C3 & C4
S1 & S3	C1 & C5
S1 & S2	C4 & C5
S1 & S2	C2 & C4

Table 2. Dimensions of the metallic (diameter) spheres and cylinders (diameter times length) used in the shape discrimination test.

Solid Aluminum Targets

Spheres		Cylinders	
SS3	7.6 cm	CS3	7.6 x 7.6 cm
SS5	12.5 cm	CS5	12.5 x 12.5 cm

Stainless Steel Water-filled Targets

Spheres		Cylinders	
SW3	7.6 cm	CW3	7.6 x 7.6 cm
SW5	12.5 cm	CW5	12.5 x 12.5 cm

Table 3. Sphere versus cylinder discrimination performance results with the foam targets. The windowed results refer to the sphere echoes for which the air-water surface reflected components in the echoes were eliminated.

256 Trials/Subject Task		Four-Subject Average (percent)	Percent Correct Windowed Average (percent)
S2	C4	96	88
S2/S3	C3/C4	93	85
S1/S3	C1/C5	88	81
S1/S2	C4/C5	84	--
S1/S2	C2/C4	91	83

Table 4. Sphere versus cylinder discrimination performance results with the metallic targets.

Percent Correct

256 Trials/Subject Task		Two-Subject Total (percent)
Solid Aluminum Targets		
SS3	CS3	100
SS5	CS5	99
SS3 & SS5	CS3 & CS5	99
Stainless Steel Water-Filled Targets		
SW3	CW3	95
SW5	CW5	99
SW3 & SW5	CW3 & CW5	94

Figures 4 through 6 show echoes from foam, solid aluminum, and hollow steel spheres and cylinders. The cylinder echoes contain slightly more energy at high frequencies. The shape discriminations were the only tasks in which spectral rather than temporal cues were dominant. The subjects used a pitch cue to make sphere-cylinder discriminations even when echoes resulted from different types of targets.

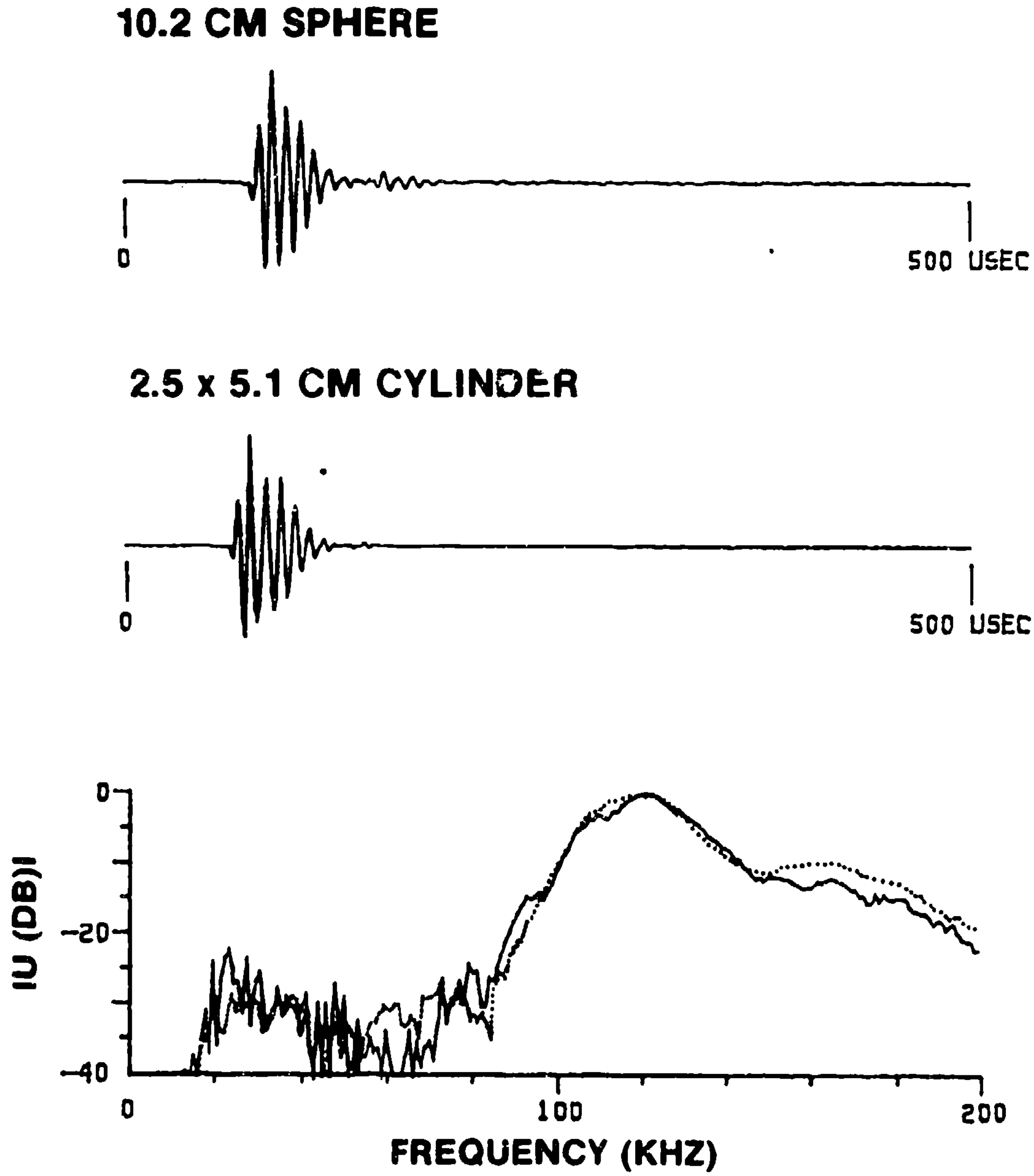

Figure 4. Typical echo waveforms and frequency spectra for solid aluminum sphere and cylinder for the shape discrimination test. The solid spectrum is for the sphere and the dotted spectrum is for the cylinder. The dimensions are the diameter for the sphere and diameter times length for the cylinder.

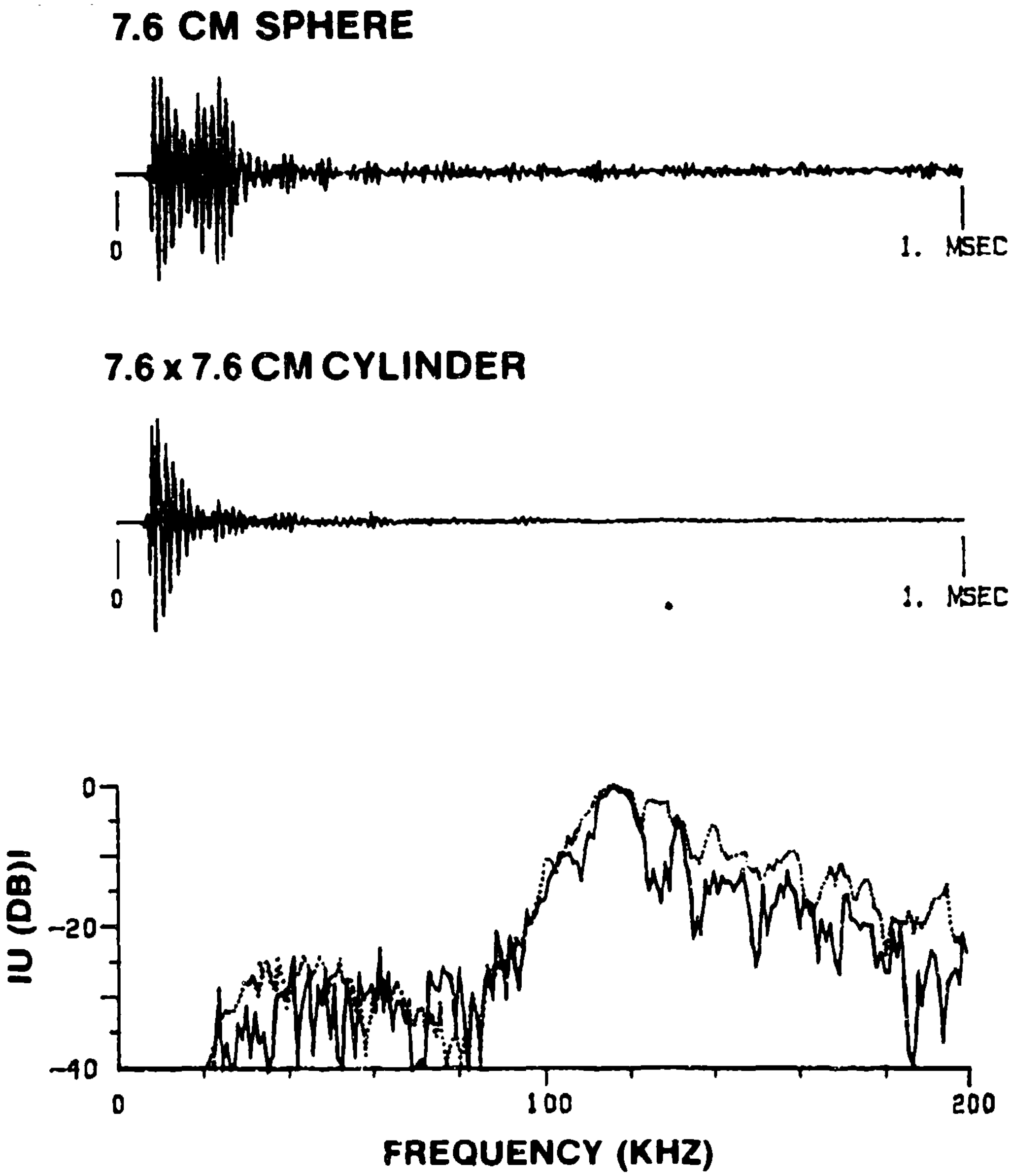

Figure 5. Typical echo waveforms and frequency spectra for foam sphere and cylinder used in the shape discrimination test. The solid spectrum is for the sphere and the dotted spectrum is for the cylinder. The dimensions are the diameter for the sphere and the diameter times length for the cylinder.

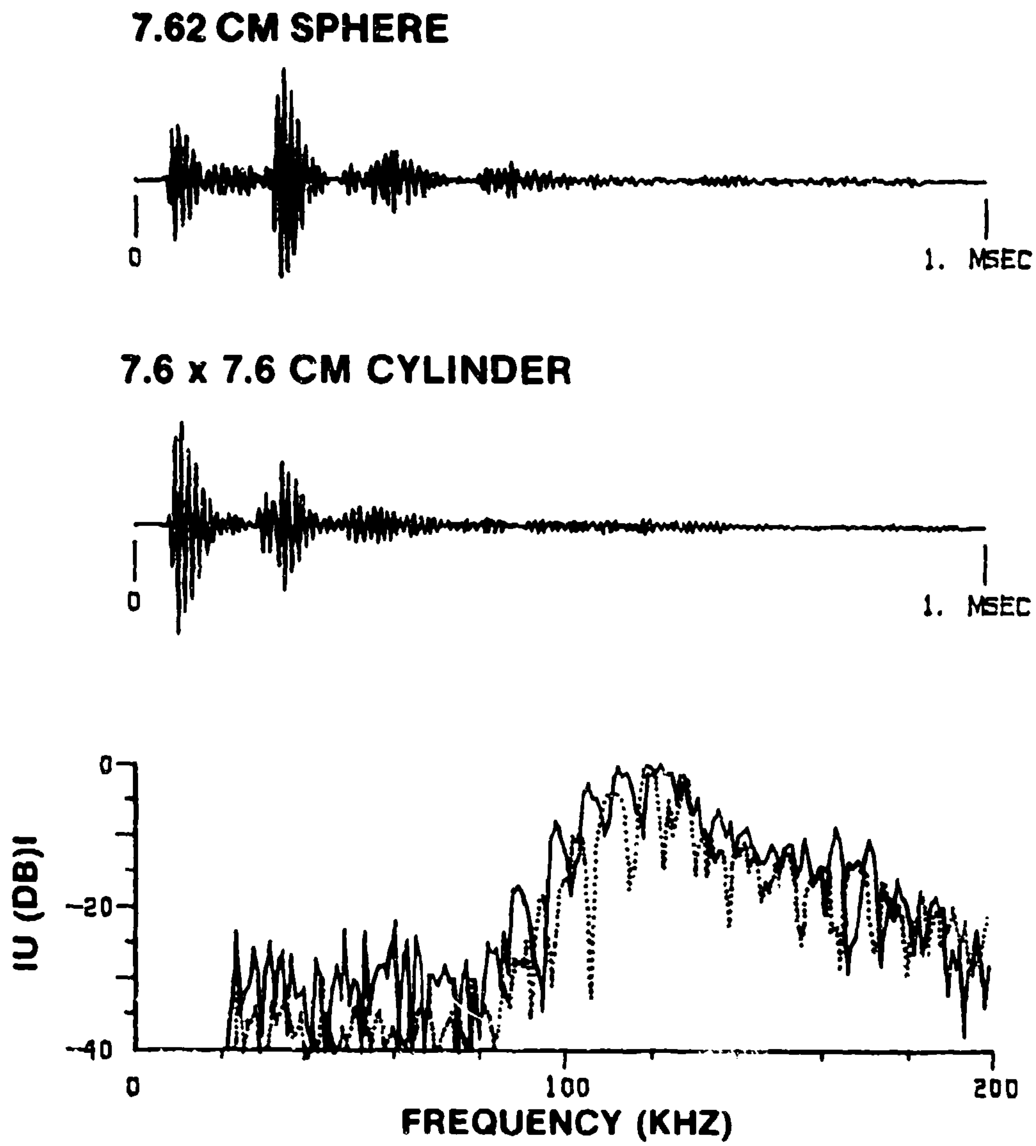

Figure 6. Typical echo waveforms and frequency spectra for water-filled stainless steel sphere and cylinder used in the shape discrimination test. The solid spectrum is for the sphere and the dotted spectrum is for the cylinder. The dimensions are the diameter for the sphere and the diameter times length for the cylinder.

EXPERIMENT III: ASPECT-INDEPENDENT TARGET DISCRIMINATION

This experiment tested whether subjects could learn to discriminate between pairs of targets, differing in material composition or internal structure, independent of target aspect. Previous research (reference 37) showed that echo waveforms for cylinders changed dramatically with aspect changes as small as 2 degrees. Our experiments determined whether subjects could generalize target discrimination cues to unlearned aspects after training at 0, 45, and 90 degrees. Five targets were used. Each target was 17.6 cm in-diameter and 17.1 cm long. Targets included solid coral rock, solid aluminum, air-filled aluminum, water-filled aluminum, and water-filled steel cylinders. Broadside aspect was defined as 0 degrees and end-on aspect as 90 degrees.

In the experiments, the subjects were first trained to discriminate between pairs of targets presented at single aspect, 0, 45, or 90 degrees. Performance exceeded 95-percent correct on each of the three tasks. Subjects were then given two-alternative forced choice discrimination tasks in which the targets could occur at any of the three previously learned aspects. Performance on these tasks was initially well below 100-percent but improved to near 100-percent correct after three to five sessions. The purpose of these intermediate tasks was to obtain baseline performance with familiar stimuli and with the subjects required to group multiple echo types into a single target class. That is, echoes from 0, 45, and 90 degrees were all mapped into the same response.

Next, discrimination was tested with echoes from targets at any of seven aspects: 0, 15, 30, 45, 60, 75, or 90 degrees. Subjects were again required to categorize each echo as targets A or B. Initial classification performance on echoes from the four new aspects measured how well training at 0, 45, and 90 degrees could be generalized to the unlearned echoes.

RESULTS AND DISCUSSION

Results of aspect-independent cylinder discrimination tests are in figure 7 and tables 5 and 6. Figure 7 shows performance for target aspects of 15, 30, 60, and 75 degrees, over time, for the discrimination between rock and water-filled aluminum targets. Performance during the initial session indicated the subject's ability to generalize previously learned cues to the new aspects. Performance was mediated by correct-response feedback after each trial. Except for the aluminum-steel discriminations, performance on all tasks was similar to that shown in figure 7. With aluminum-steel discriminations, subjects could not generalize cues to the new aspects, and the improvement in performance over time was small.

Table 5 shows average performance for the four new aspects with data pooled across subjects for each discrimination task. The table shows performance for the first and for the last three sessions. Initial transfer of learning resulted in 72- to 80-percent correct discriminations for echoes at the four new aspects, with performance improving to 90-percent correct after 14 sessions. Initial performance for aluminum versus steel cylinders was chance level, but improved to 76-percent correct after 14 sessions.

Table 5. Aspect-independent cylinder discrimination, 15, 30, 60, and 75 degrees. Data are pooled across subjects.

Task	Session 1	Sessions 12-14
Water-filled aluminum vs steel	55.8	76.0
Water-filled vs air-filled aluminum	72.0	91.5
Water-filled aluminum vs solid rock	79.9	95.1
Water-filled vs solid aluminum	73.0	83.5
Solid aluminum vs rock	79.6	89.5

Table 6. Overall material and internal structure discrimination performance results for the cylindrical targets at the different aspect angles. The aspect angle is the angle between the direction of the incident signal and a normal to the longitudinal axis of the cylinder.

Degrees	Hollow Alum	Coral Rock	Hollow Alum	Solid Alum	Solid Alum	Coral Rock	Hollow Alum	Hollow Steel	Alum Water	Alum Air
0	89	92	56	93	93	94	93	97	90	100
15	98	93	99	76	88	91	97	57	100	80
30	89	93	77	73	94	92	53	91	69	91
45	88	98	83	93	81	93	84	87	95	91
60	81	85	81	90	80	87	53	69	71	92
75	98	80	91	58	95	70	68	71	83	98
90	97	96	97	98	97	92	100	95	99	98

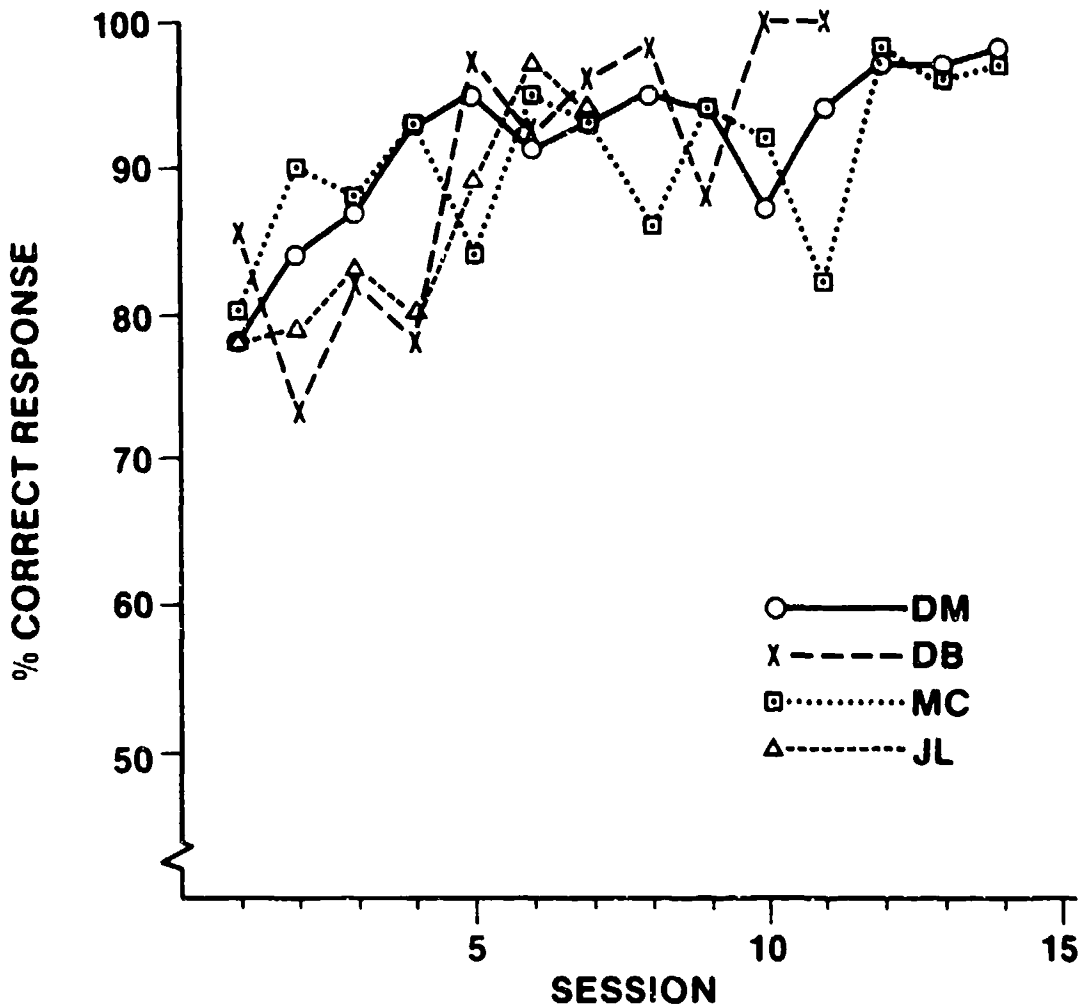

Figure 7. Performance for discrimination between hollow aluminum and solid rock cylinders at novel aspects of 15, 30, 60, and 75 degrees as a function of session number.

Table 6 shows response accuracy as a function of target aspect for each discrimination. Data are pooled across subjects. Although performance at 0, 45 and 90 degrees was initially high, several entries in table 6 represent poor performance for these echoes. Subjects were asked to group echoes from seven different aspects into a single category for each target. Such instructions required subjects to remember general cues, at the expense of specific cues of a particular echo. Thus, errors on some targets of the training set were not surprising.

Subjects reported that echoes from targets at different aspects sounded very different. They used many different cues for the discriminations and reported that training with the 45-degree echoes provided the most useful cues. Figure 8 shows echo waveforms for hollow and solid aluminum cylinders at 0, 45, and 90 degrees. The differences are obvious. For cylinder echoes other than at 0 and 90 degrees aspect, the first echo component, representing a front surface reflection, is generally not the largest. This fact suggests the possibility of salient rise time cues for discrimination, although such cues were not specifically reported by subjects. We cannot identify cues that subjects might use to discriminate between the echoes in figure 8, independent of aspect. However, the data show that subjects receiving training at a few widely separated aspects can generalize discrimination cues to previously unlearned aspects.

EXPERIMENT IV: DETECTION AND DISCRIMINATION IN NOISE

Broadband sonar echo detection and discrimination experiments with human subjects were conducted in white noise. The measurement of human performance in noise can answer fundamental questions. For example, what is the difference in signal-to-noise (S/N) ratios between the point where echoes are just detectable and the point where they can be discriminated? This information is a direct measure of task difficulty and can give insights into the importance of particular discrimination cues. If part of a signal that is 20 dB below the peak is required to discriminate it from a similar signal, the discrimination threshold must be at least 20 dB higher than the detection threshold.

Performance data for subjects' detection and discrimination are presented as psychometric functions, which show the probability of a correct response as a function of S/N ratio. (S/N is defined here as the ratio of total integrated signal energy to noise power in a 1-Hz band.)

METHODS-DETECTION OF SPHERE ECHOES

Psychometric functions were obtained for two subjects detecting echoes from a 3-inch water-filled, stainless-steel sphere in white noise. In two experiments, detectability was measured as a function of time-expansion factor and echo-repetition rate using a modified method of constants. The time-expansion factor is the ratio of input to output sample rate for the echoes and has a value of 50 for the experiments reported so far. Echoes were time-expanded so their center frequencies were in the human audio range. If the time-expansion factor is changed, both the center frequency and the duration of an echo also change.

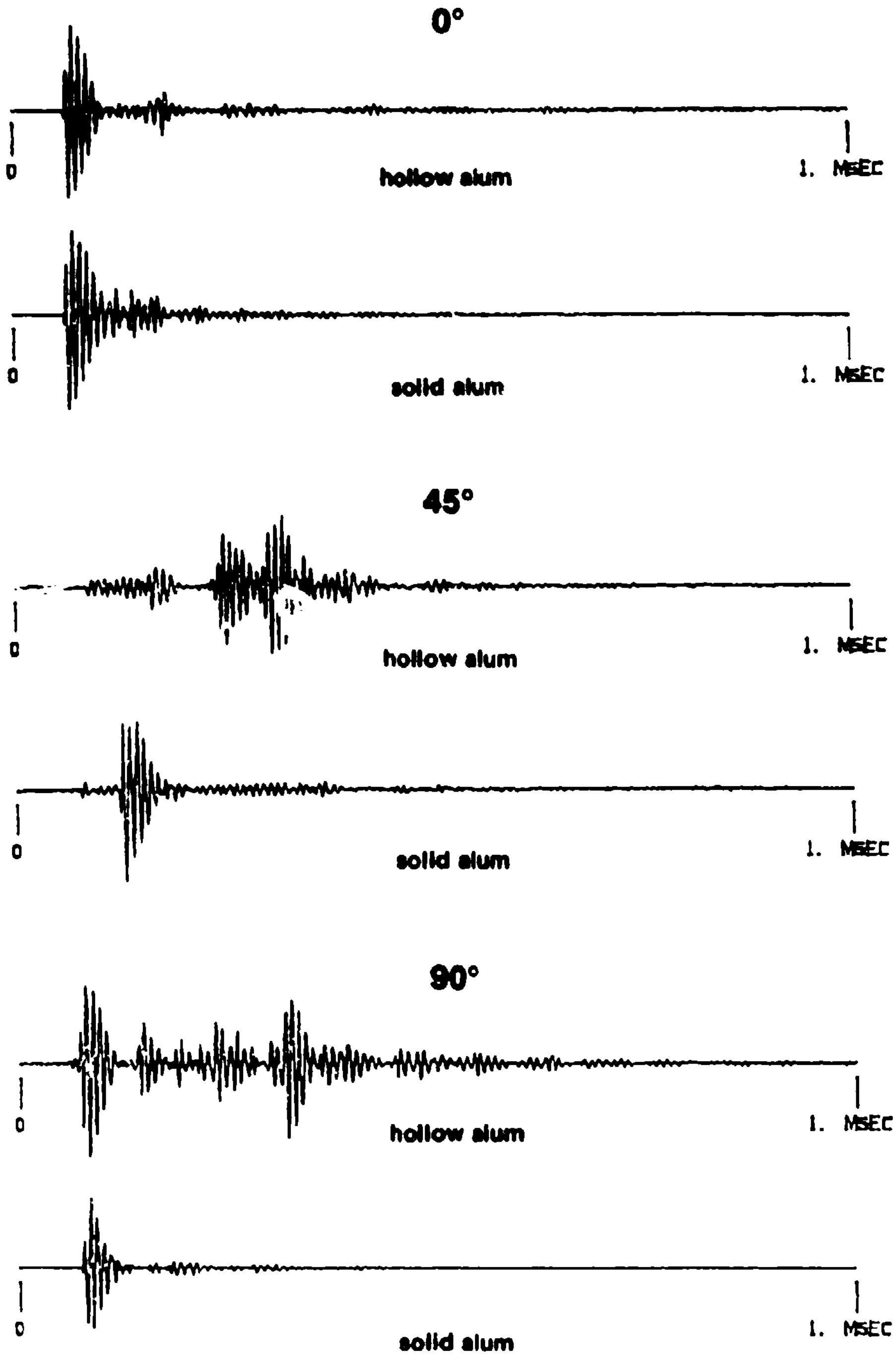

Figure 8. Typical echo waveforms for the hollow and solid aluminum cylinders at the baseline aspects of 0, 45, and 90 degrees.

The detectability of sphere echoes was measured in 100 trial sessions. Subjects reported whether a given trial contained an echo plus noise, or noise alone. The S/N ratio was constant for each block of 10 trials. Five S/N ratios were tested per session in 10 blocks randomly presented. Two S/N ratios at each level made a session, and the first block of trials was always run at the highest S/N ratio. Au and Penner (reference 38) used the same procedure to measure sphere detection by dolphins.

In the first experiment, the repetition rate was 32 echoes per second, the same repetition rate used by the dolphin (reference 38). The time-expansion factor, with values of 25, 50, or 75, was varied randomly between sessions. Center frequencies and durations for the three time expansions are as follows:

Time Expansion	Center Frequency	Signal Duration
25	4.8 kHz	10 ms
50	2.4 kHz	20 ms
75	1.6 kHz	30 ms

Because the echoes result from digital-to-analog conversion, doubling signal duration is equivalent to doubling the signal energy. This is taken into account in the definition of S/N ratio given above.

In experiment ..., the time-expansion factor was 50, and the echo-repetition rate was varied between 16, 24, and 32 pulses per second randomly for different sessions. This experiment tested how echo detectability might change as a function of duty cycle, given a constant echo center frequency. (Duty cycle here refers to the percentage of "on time" for the echoes.)

RESULTS AND DISCUSSION

Figure 9 shows results from one subject at the three time-expansion factors of experiment 1. At each S/N ratio, 140 trials per subject were collected. The 75-percent correct response thresholds for subject DM were 5.5, 8.0, and 11.5 dB for time expansions of 75, 50, and 25, respectively (figure 9). Thresholds for subject RB were 6.0, 6.3, and 9.1 dB for the same time-expansion factors. Thus, echoes were hardest to detect at 4.8 kHz center frequency, with a time-expansion factor of 25. Several factors could explain this result. In the frequency region from 2.4 to 4.8 kHz, the critical masking bandwidth of the ear is more than doubled for a doubling of signal bandwidth resulting from different time expansions. Thus, more noise contributes to masking for a time-expansion factor of 25 than for a time-expansion factor of 50. Also, the detectability of the echoes may be affected because the signal duration and, therefore, its duty cycle is a function of the time-expansion factor.

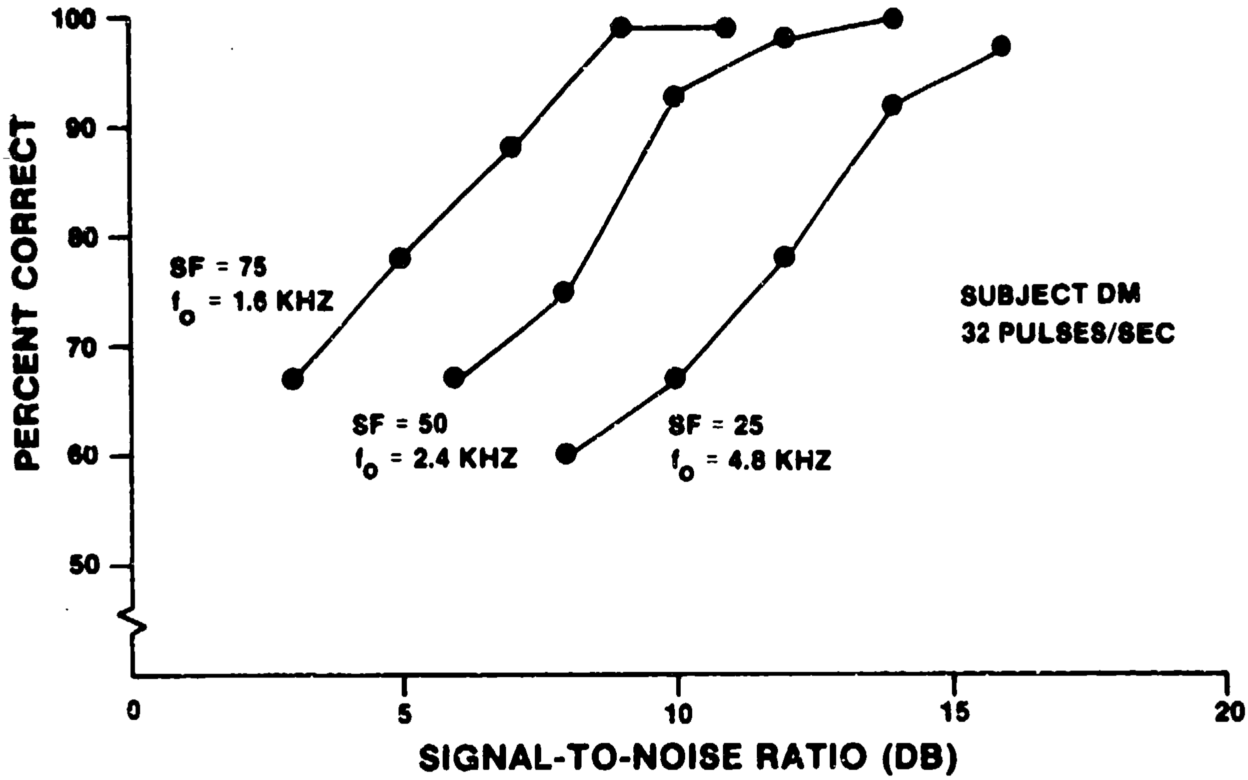

Figure 9. Sphere detection in noise performance for different stretch factors (SF). The SF and peak frequencies are listed with each performance curve.

The second experiment tested whether differences in detectability, noted in Experiment I, are related to changes in echo center frequency or t› echo duty cycle. For repetition rates of 16, 24, and 32, the duty cycles were 1/3, 1/2, and 2/3, respectively. The 75-percent detection threshold at a repetition rate of 16 echoes/sec was 1 dB higher than the threshold at 32 echoes/sec. The threshold value for a repetition rate of 24 was between those of 16 and 32. Three subjects produced the same results. Absolute thresholds differed between subjects, yet the performance difference between the highest and lowest repetition rates was always about 1 dB and was never statistically significant. Thus, the effect of changing the duty cycle or the fraction of the on time from 2/3, to 1/2, to 1/3 is small, and it does not account for the performance in experiment 1. The detectability differences observed in experiment 1 probably resulted from masking factors related to the critical bandwidth (reference 39). The results of experiment II show that changes in repetition rate did not significantly affect detection performance over the range tested. This result cannot be generalized to echo detection at slow repetition rates, e.g., the cylinder detection experiment conducted at 4 pulses per second, which is discussed next.

METHODS-DISCRIMINATION OF ‹ ‹LINDER ECHOES

In the following experiments, the echo time-expansion factor was 50, and the repetition rate was 4 pulses per second. The psychometric functions were obtained with a modified method of constants with 10-trial blocks.

Prior to measurements for material composition discrimination, baseline detection thresholds were determined for three subjects using echoes from 3.8- and 7.6-cm-diameter, water-filled aluminum cylinders. These cylinders were used in the previous material composition experiments. Pooled across subjects, the 75-percent thresholds were 10.5 dB for the large cylinder echoes and 10.7 dB for the small cylinder echoes. In later target discrimination experiments, S/N ratios were given with respect to these detection thresholds. Detection thresholds were similar for the small and large aluminum cylinders (the peak amplitudes were normalized), so we assumed that the echoes from different cylinders were equally detectable.

The minimum requirement for discrimination between stimuli is the detection of a feature that is different between the stimuli. If two stimuli are completely dissimilar, (i.e., having no common features), the discrimination problem reduces to a detection problem, and the discrimination and detection thresholds will be equal (reference 40). As the similarity between stimuli increases, the difference in S/N ratios between the discrimination and detection thresholds also increases.

The results for material composition discrimination tasks are shown in figures 10, 11. The slopes of these functions are much less steep than the slopes of the detection functions of figure 9. These gradual slopes imply the use of multiple discrimination cues, since noise masking of a single feature does not completely degrade discrimination performance. Subjects probably increased their reliance on secondary and less effective cues as the S/N ratio was decreased.

Differences between the discrimination and detection thresholds for four target pairs are listed in table 7. Simple tasks such as discrimination between aluminum-bronze cylinders or between solid-hollow cylinders require S/N ratios 7 to 10 dB above the detection threshold to obtain 75-percent discrimination. The most difficult material discrimination, 7.6-cm-diameter aluminum versus glass cylinders, requires a S/N ratio 25 to 30 dB above the detection threshold for 75-percent discrimination. In this task, subjects used cues that were at least 25 dB below the peak amplitudes of the signals.

Table 7. Difference in S/N ratio between the 75-percent detection and discrimination thresholds for four tasks. An average detection threshold of 10.5 dB was used for all cylinders.

Task	DM	PT	RB
Solid vs water-filled aluminum, 7.62 cm	13.7	13.8	----
Water-filled aluminum vs glass 7.62 cm OD	23.9	29.5	20.9
Water-filled aluminum vs glass 3.81 cm OD	10.5	----	----
Water-filled aluminum vs bronze 3.81 cm OD	7.0	----	11.2

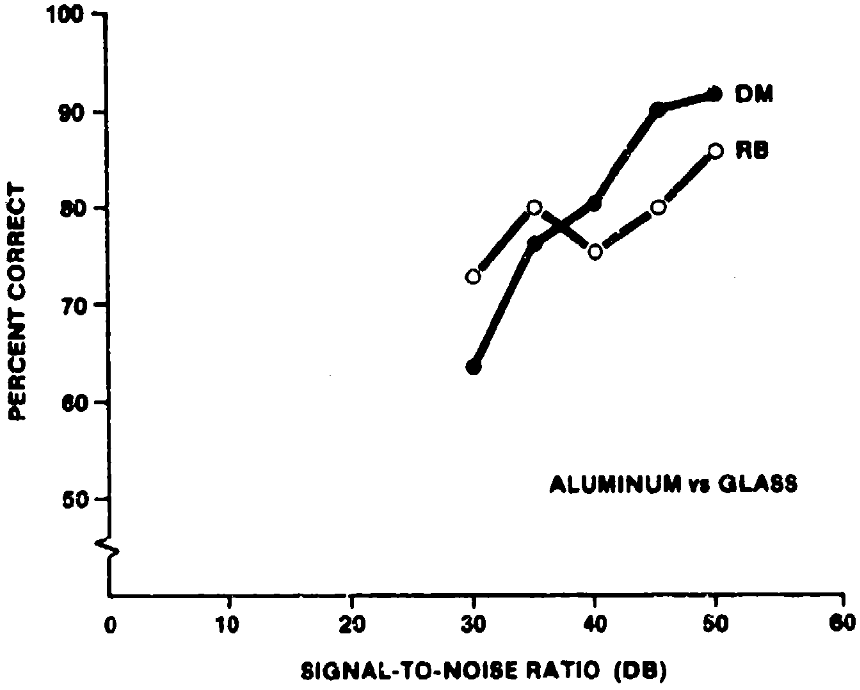

Figure 10. Material composition discrimination performance results in noise between the hollow aluminum and glass cylinders.

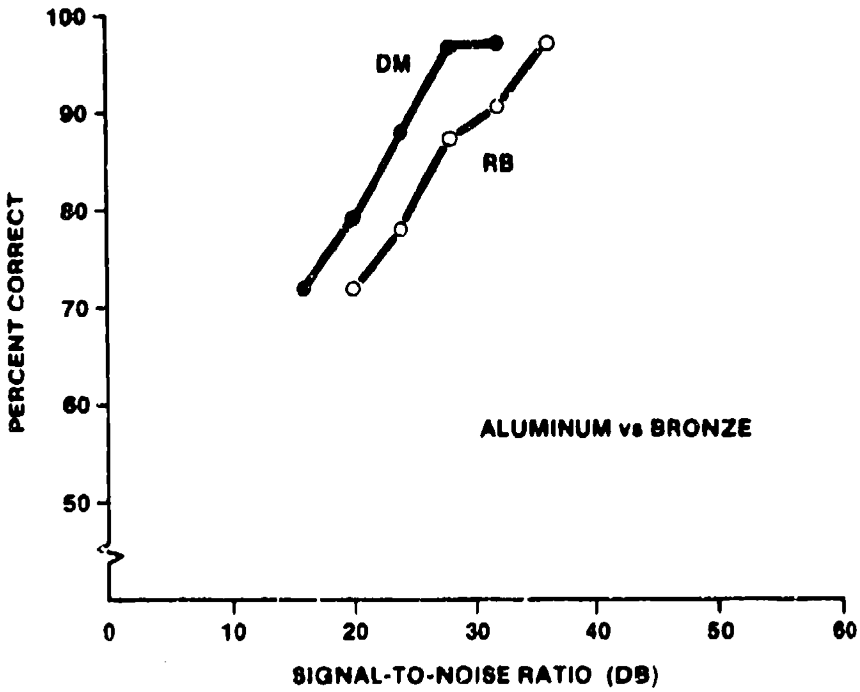

Figure 11. Material composition discrimination performance results in noise between the hollow aluminum and bronze cylinders.

FEATURE EXTRACTION AND PATTERN RECOGNITION

Software was developed to extract acoustic features from target echoes, similar to the features identified by subjects in our discrimination tests. Because highlight separation and highlight amplitude ratios are necessary determinants of both time-separation pitch and echo duration, the software extracted these time-domain features from the signal envelopes. Echoes were synthesized from the extracted feature sets, and usually the synthetic echoes contained the same discrimination cues as the real echoes. A notable exception involved sphere-cylinder discrimination. Here, the synthetic echoes could not be accurately discriminated because the feature extraction/synthesis process did not preserve the necessary spectral cues. Feature subsets were also used to synthesize signals for discrimination tests to determine the relative importance of the features. Two types of automatic pattern recognition algorithms were tested. The first used the time-domain features described by subjects in the discrimination tests of the previous section. The second, a filter bank model developed by Chestnut and Floyd (reference 37), used spectral features. Results of the feature extraction, echo synthesis, and automatic pattern recognition experiments are presented in this section.

The software that extracted time-domain features was a highlight detector. The software measured probability of occurrence, time separation from the largest highlight, and amplitude ratio relative to the largest highlight for each highlight in a group of echoes. The first stage of the processor was a peak detector that stored information about every point in a signal where the slope changed from positive to negative. A series of criteria selected the extrema, which were defined as highlights. Small-amplitude extrema in the immediate neighborhood of a larger maximum were rejected. After obtaining a list of highlights for a given signal, the absolute maximum was assigned a time separation of 0 and an amplitude ratio of 1. The other highlights were assigned negative or positive time separations according to position before or after the absolute maximum. Amplitude ratios for each highlight were calculated with respect to the maximum. If the reflection with the largest amplitude occurred first, all highlights had positive time separations.

In the second stage, the software aligned the features across signals. Absolute maxima were aligned so that every group of signals had a highlight occurring with probability 1.0, amplitude ratio 1.0, and time separation 0. Other highlights were aligned, and amplitude ratios and time separations became statistical quantities represented as means and standard deviations for the group of signals. The probability of occurrence for each highlight was calculated; i.e., those that occurred in only one signal had probability $1/n$, where n was the number of signals in the input. Means and standard deviations for both time separation and amplitude ratio were also calculated. The software also calculated statistics for highlight rise times for some types of input signals. Input signals could be unprocessed echoes, echo envelopes, envelopes of matched filter responses, etc. Measures of rise time were not defined for raw echoes, which contain many zero crossings. Feature lists statistically defined each highlight in a group of signals.

The next stage of software processing determined whether the extracted features defined separable target classes. The features were the input to an automatic pattern recognition algorithm (discussed later) and were also used to create synthetic target echoes for human discrimination.

The synthetic echoes served two purposes. First, discrimination tests with these echoes determined whether the features described separable target classes. Human listeners indicated whether the synthetic echoes sounded similar to real echoes, so we could determine whether the feature extraction process preserved the appropriate discrimination cues. Second, subsets of the feature list were used to create synthetic signals for discrimination, so we could determine the relative importance of individual features. For example, if echoes from two targets are synthesized using only two highlights from each feature list, the saliency of time-separation pitch can be measured for this discrimination.

SYNTHETIC ECHOES

Echo envelopes were used as input for the feature xtraction process, which produced the feature lists used to make synthetic echoes. Echo envelopes provided more reliable feature lists than unprocessed echoes. Highlights in the synthetic echoes were created using portions of the incident signal. The position and amplitude of each highlight was a Gaussian random variable, with means and standard deviations taken from the feature list. The probability of occurrence for each highlight was also determined from the feature list. Because synthetic highlights were replicas of the incident signal, rise times and spectral properties were identical for all highlights. Thus, the synthetic echoes did not contain information about frequency-dependent reflection characteristics of the targets. Only time-domain information involving highlight separation and amplitude was preserved. Figure 12 shows real and synthetic echoes for a solid aluminum cylinder, and figure 13 shows a synthetic echo and an edited synthetic echo composed of the two highest amplitude highlights. While minor differences existed in the fine structure of the echoes, the overall echo similarity is obvious.

In all tests, human subjects easily discriminated synthetic echoes from their real counterparts using differences in rise time and spectral characteristics. However, except for sphere-cylinder discriminations where spectral information is critical, subjects reported using the same cues in synthetic echo discriminations as in the original discriminations. Thus, synthetic echoes from solid aluminum cylinders produced metallic ringing aluminum-glass discriminations were still based on a duration cue, and TSP cues still described differences between many synthetic target echoes. Direct transfer of learning from real to synthetic echoes was not measured.

While the complete synthetic echoes described distinct target classes, discrimination tests conducted with edited synthetic echoes were inconclusive. The quality of such sounds differed markedly from that of the complete echoes, so we could not determine what cues remained relevant. However, we found that echo subsets composed of only two highlights could be discriminated using time-separation pitch for those cases where original performance had been attributed to TSP. The perceived pitches for the subsets were different from the pitches of the original echoes.

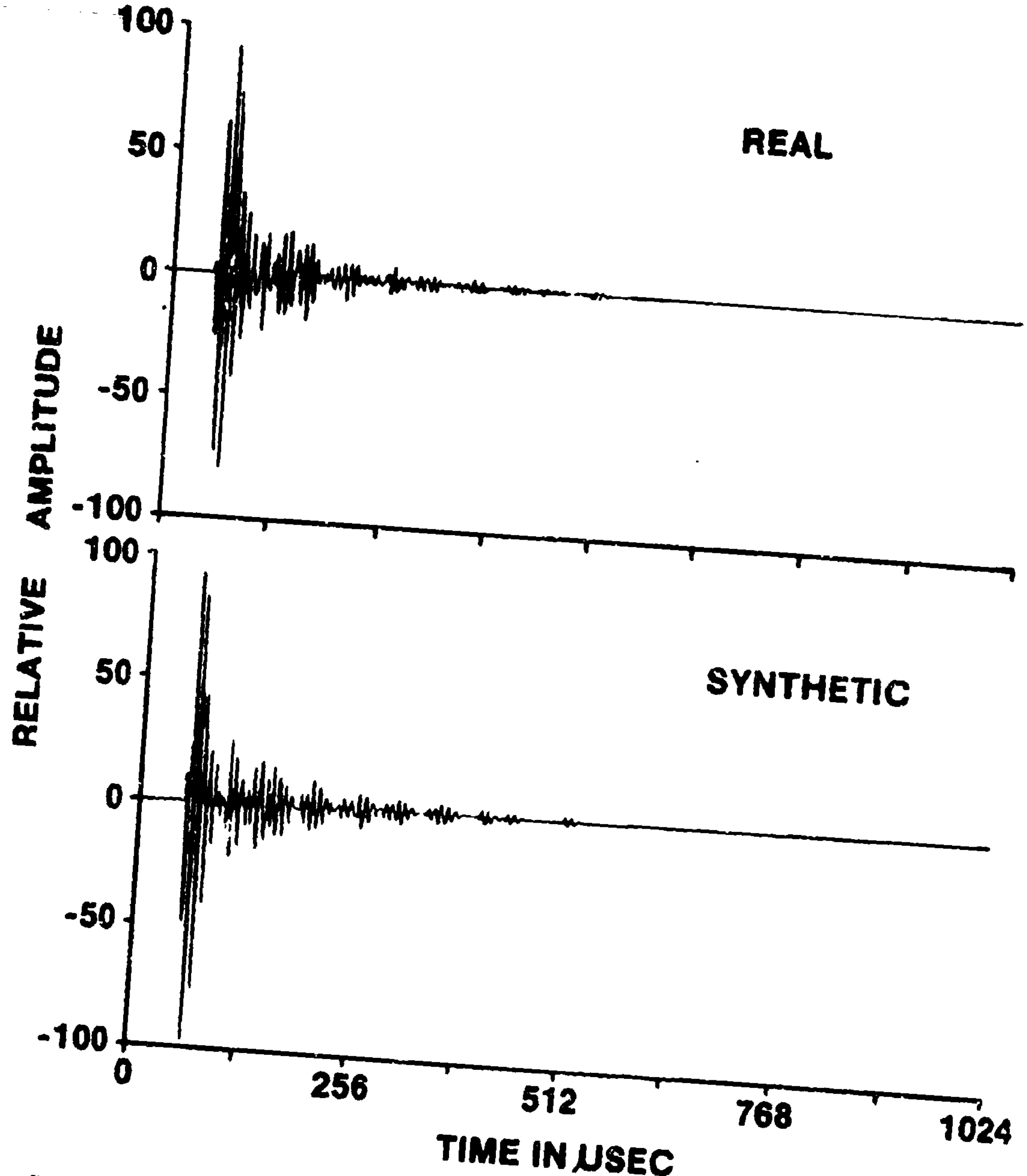

Figure 12. Examples of a real echo and a synthetic echo for a solid aluminum cylinder.

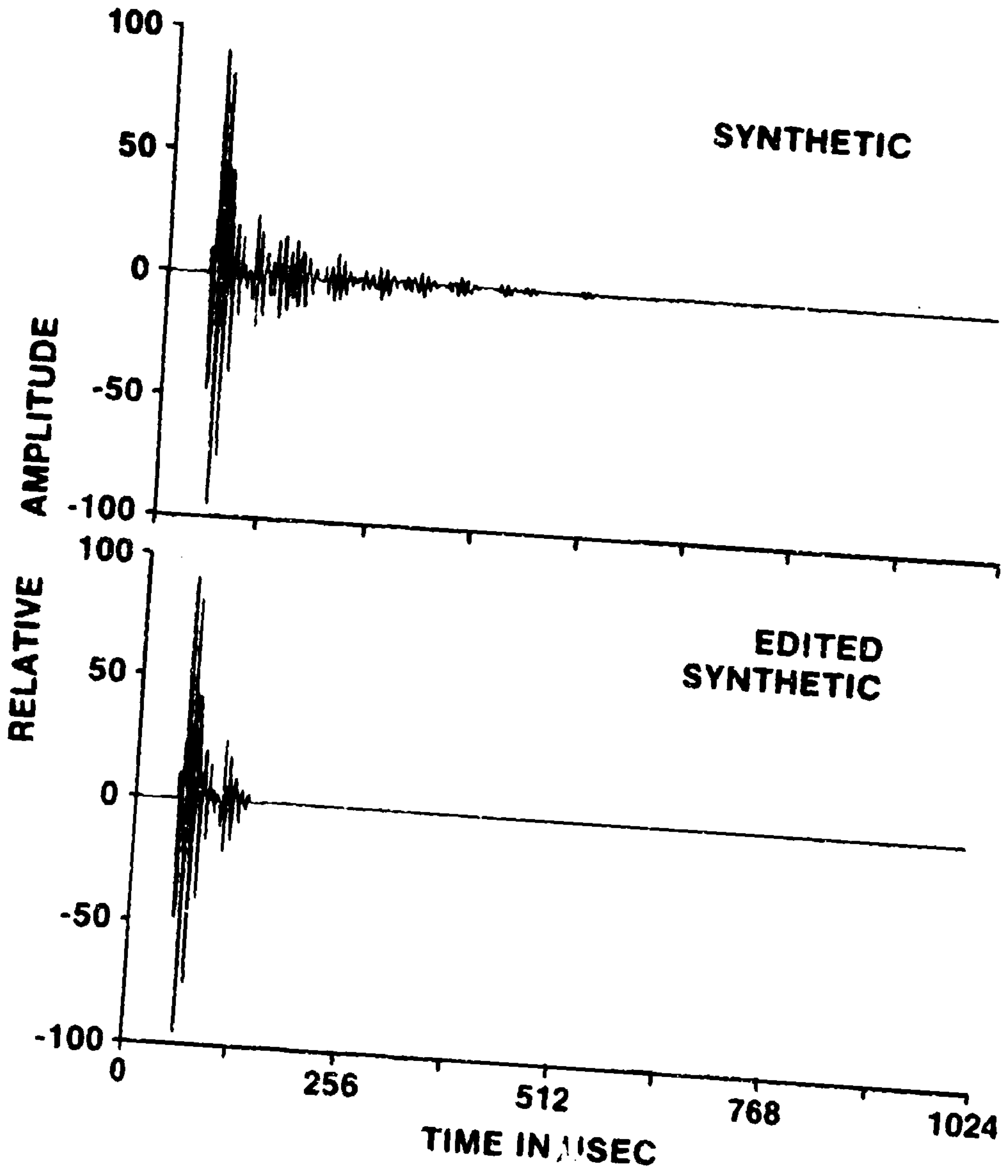

Figure 13. Examples of a synthetic and an edited synthetic echo for a solid aluminum cylinder.

PATTERN RECOGNITION

The accuracy of the extracted feature sets was tested using automatic pattern recognition algorithms. The pattern recognition algorithms were part of the Interactive Laboratory System (ILS) software package developed by Signal Technology Inc. (reference 41). The algorithms classified target echoes by calculating the Euclidian distances between sets of test and reference feature vectors. Reference and test data were represented by vectors of 25 features; each feature represented the time-separation and relative amplitude of an echo highlight.Reference data were means from 10 echoes. Feature vectors represented highlight amplitude as a function of time, with both variables calculated relative to parameters of the largest highlight. The time axis was partitioned into bins of 20 points each, and each bin was assigned an amplitude value determined by highlights in that portion of the echo. A time resolution of 20 points in each partition is equivalent to 20 microseconds of time separation in the original echoes. Because each feature vector contained 25 elements, time separations as large as 500 microseconds (25 x 20) were represented. When any of the 25 time windows contained an echo highlight, that element was assigned the value of the highlight amplitude ratio relative to the maximum. If the partition did not contain a highlight, a value of 0 was assigned. When more than one highlight was present in a partition, the largest amplitude was used. Within the constraint of a 20-point time resolution, the Euclidian distance between two vectors constructed in this way is a measure of stimulus similarity in terms of the extracted feature sets.

Performance of the feature extraction and pattern recognition algorithms for material composition discrimination are shown in tables 8 and 9 as confusion matrices. Rows of the matrices represent test echoes; columns represent reference echoes (mean vectors). The algorithm calculated the distances from each test echo to each reference vector and identified each with the minimum distance reference class. Elements on the matrix diagonal represent correct responses; off-diagonal elements represent confusions. The algorithm's performance was 90-percent correct for material composition (table 8) and 100-percent correct for internal structure (table 9). Chance performance was 14.3-percent correct (1 in 7) for material composition and 25-percent correct (1 in 4) for internal structure. When echoes from two different glass cylinders were added to the data of table 8, to make a 9- by 9-confusion matrix, performance dropped to 62-percent correct; chance dropped to 11-percent correct (1 in 9). Many glass test echoes were incorrectly identified, and some echoes from other cylinders were wrongly identified as glass. Identification of echoes from the glass cylinders was also poor for both humans and dolphins.

The test echoes were originally collected in several recording sessions for different purposes. For example, the 38 echoes from the 7.62-cm-diameter aluminum cylinder were represented to the pattern recognition software as follows: 10 reference echoes originally used for material discrimination experiments, 9 test echoes originally used for the 0-degree signals in the aspect-independence discrimination study, and 19 test echoes collected about 1 year later, the transducer having been removed and replaced in the water. Subjects reported that echoes from each target sounded similar, regardless of the echo's collection set.

The time-domain algorithm's performance was not degraded by using stimuli collected at different times. This was, however, not the case for a spectral feature extraction algorithm (reference 6) to be discussed next. That algorithm failed completely when test and reference echoes were from different recording sessions.

A spectral feature extraction algorithm used by Chestnut et al. (references 6 and 37) tested target recognition performance. The model consists of a bank of parallel and constant-Q filters, and the extracted features are samples of the target's frequency response. As above, the algorithm calculates the Euclidian distance between test and reference feature vectors and identifies test echoes with the minimum-distance reference class. Prior to feature extraction, the echo spectra are normalized by the spectrum of the incident signal. The resulting signals represent the targets' frequency-dependent reflection characteristics.

This filter bank model was tested with material composition and sphere-cylinder discriminations. Thirty constant-Q filters over the frequency range of 50 to 200 kHz were used. This frequency range corresponds to a range of 1 to 4 kHz for the time-expanded echoes to which human subjects listened.

When test and reference echoes were measured on the same day, performance of the filter bank model was 90- to 100-percent correct for material composition discriminations using the same targets as above. Discrimination between metal spheres and cylinders was 90-percent correct. The targets were the same used in the human studies, and echoes were collected at the same time.

The algorithm was then tested with the echoes from the time-domain material composition test (table 8). The spectral feature algorithm incorrectly classified all the echoes when comparisons involved echoes recorded in different sessions. Thus, the algorithm performed at 0-percent correct for the aluminum cylinder given in the earlier example, and the time-domain model, derived from studies of human discrimination, was superior to the spectral processing model.

Table 8. Confusion matrix for cylinder material composition discrimination using ILS pattern recognition software.

Reference Target Echoes

Test Target Echoes	Alum Cycl-1	Steel Cycl-1	Bronze Cycl-1 (percent)	Alum Cycl-2	Steel Cycl-2	Bronze Cycl-2	Solid Alum Cycl
Alum Cycl-1	97					3	
Steel Cycl-1		100					
Bronze Cycl-1		100					
Alum Cycl-2				93		7	
Steel Cycl-2				37	63		
Bronze Cycl-2						100	
Solid Alum Cycl					11		89

Table 9. Confusion matrix for internal structure discrimination using the ILS pattern recognition software.

Reference Test Target

Test Target Echoes	Alum Cycl-2 Airfilled	Coral Cycl-2 Solid (percent)	Alum Cycl-2 Waterfilled	Alum Cycl Solid
Alum Cycl-2 Airfilled	100			
Coral Cycl-2 Solid		100		
Alum Cycl-2 Waterfilled			100	
Alum Cycl-2 Solid				100

SUMMARY AND CONCLUSIONS

The human auditory system has excellent pattern recognition capabilities, which can be used to identify acoustic cues in broadband sonar classification tasks. Echoes from targets were collected in a test pool using a broadband ultrasonic pulse with a 120-kHz center frequency and 39-kHz bandwidth. The pulse was similar to dolphin echolocation pulses. Human subjects listened to the echoes played at one fiftieth of the original sample rate during two alternative forced choice target discrimination tests. Echo waveforms contained highlights from multiple internal reflections, with differences in highlight arrival times determined by acoustic path length differences in the targets.

Subjects' discrimination of material composition exceeded 95-percent correct for water-filled target cylinders of aluminum, bronze, and steel. Differences in time-separation pitch associated with correlated echo highlights were the primary cues used by the subjects. Discrimination between aluminum and glass cylinders was more difficult, with differences in echo duration as cues.

Sphere-cylinder discriminations using foam, solid aluminum, and water-filled steel targets were all above 85-percent correct. The same discrimination cues were used for all target types, the most salient being a higher spectral pitch for cylinder echoes and low-frequency reverberation in the sphere echoes. The sphere-cylinder discrimination was the only task in which subjects used spectral and not temporal cues.

Subjects were trained to discriminate between cylinders differing in material and internal structure at 0, 45, and 90 degrees. Transfer of learning was measured with new random-aspect echoes in 15-degree increments. Generally, subjects attained over 75-percent correct performance with the new echoes. They reported that training at 45 degrees provided the most cues.

Discrimination tests in noise showed that simple tasks using solid and hollow aluminum cylinders required S/N ratios about 10 dB above the detection threshold for 75-percent correct discrimination. Difficult tasks such as aluminum versus glass cylinder discrimination required a 30-dB difference between the 75-percent detection and discrimination thresholds. Psychometric functions were not steep, implying that subjects used secondary discrimination cues when noise masked the primary cues. In all the discrimination tasks tested, subjects' learning was sudden and insightful rather than gradual.

Discrimination cues identified from the tests with human subjects were used to design software to extract acoustic features. Because highlight separation and highlight amplitude ratios are necessary determinants of both time-separation pitch and echo duration, these time domain features were extracted from the signal envelopes. Echoes were synthesized from the feature sets, and, generally, the synthetic echoes contained the same cues as the real echoes.

Discrimination value of the feature sets was tested in an automatic pattern recognition algorithm. The algorithm calculated Euclidian distance between mean reference vectors and feature vectors for test echoes. The algorithm achieved 90-percent accuracy on a material discrimination test with seven targets, where chance performance would have been one of seven or 14-percent correct. The results indicate

that the human auditory system can provide information useful for developing signal processing algorithms for sonar target recognition. Application of similar methods to characterize features of other targets should enhance the development of automatic recognition algorithms. Further investigation is needed to determine performance of the present algorithms in noisy and reverberant environments. In addition, since aspect-independent target discrimination was demonstrated with humans, the feature extraction software should be expanded to include this capability.

Computer-assisted classification could reduce both operator training time and operator stress associated with classification decisions. Algorithms that identify the probability of successful target classification should also reduce the time required to make a classification decision by reducing an operator's false-alarm rate. The benefits of algorithms using recognition features similar to those used by the human auditory system have been demonstrated in this study.

REFERENCES

1. Brill, D., and Uberall, H. (1971). Acoustic Waves Transmitted Through Solid, Elastic Cylinders. J. Acoust Soc Am. vol 50(3), p 921-939.

2. Peterson, B., Varadan, V.V., and Varadan, V.K. (1980). Scattering of Acoustic Waves by Layered Elastic and Viscoelastic Obstacles in Water. J Acoust Soc Am. vol 68(2), p 673-685.

3. Fish, J.F., Johnson, C.S., and Ljungblad, D.K. (1976). Sonar Target Discrimination by an Instrumented Human Diver. J Acoust Soc Am, vol 59(3), p 602-604.

4. Howard, J.H., Jr. (1977). Psychological Structure of Eight Complex Underwater Sounds. J Acoust Soc Am, vol 62(1), p 149-156.

5. Webster, J.C., Woodhead, M.M., and Carpenter, A. (1973). Perceptual Confusions Between Four-Dimensional Sounds. J Acoust Soc Am, vol 53, p 448-456.

6. Chestnut, P., Landsman, H., and Floyd, R.W. (1979). A Sonar Target Recognition Experiment, J Acoust Soc Am, vol 66(1), p 140- 147.

7. Reed, S.K. (1973). Psychological Processes in Pattern Recognition. Academic Press, New York.

8. Martin, D.W. (1978). Discrimination of Noise-like Sounds Involving Multiple Interacting Features. Pennsylvania State University Masters Thesis.

9. Gilliom, J.D., Klein, C., and Taylor, D.W. (1977). Harmonic Configuration, the Effect on the Hearing-Out Process. J Acoust Soc Am, vol 61, S-1.

10. Janota, C.P. (1977). An Experimental Treatment of Auditory Discrimination of Complex Noiselike Sounds. Pennsylvania State University Doctoral Dissertation.

11. Howard, J.H., Jr., and Silverman, E.B. (1975). Structural Components in the Perception of Sixteen Complex Sounds. Human Performance Laboratory, Catholic University of America, Technical Report 75-1.

12. Howard, J.H. Jr., and Ballas, J.A. (1978). Feature Selection in Auditory Perception. Human Performance Laboratory, Catholic University of America, Technical Report.

13. Rice, C.E. (1966). The Human Sonar System, in Animal Sonar Systems, Biology and Bionics. R.G. Busnel, ed., Laboratoire de Physiologie Acoustique INRA-CNRZ, Jouy-en-Josas-78, France.

14. Schnitzler, H.U., and Henson, O. Jr. (1980). Performance of Airborne Animal Sonar Systems: I. Microchiroptera, in Animal Sonar Systems. R.G. Busnel and J.F. Fish, ed., Plenum Press, New York and London.

15. Fenton, M.B., (1980). Adaptiveness and Ecology of Echolocation in Terrestrial Aerial) Systems, in Animal Sonar Systems. R.G. Busnel and J.F. Fish ed., Plenum Press, New York and London.

16. Basset, I.G., and Eastmond, I.J. (1964). Echolocation Measurement of Pitch versus Distance for Sounds Reflected from a Flat Surface. J Acoust Soc Am, vol 36(5), p 911-916.

17. Ammons, J.H., Worchel, P., and Dallenbach, K.M. (1953). Facial Vision: The Perception of Obstacles Out of Doors by Blindfolded and Blindfolded-Deafened Subjects. Amer Journ Psych, vol 66, p 519-553.

18. Cotzin M., and Dallenbach, K.M. (1950). Facial Vision, the Role of Pitch and Loudness in the Perception of Obstacles by the Blind. Amer Journ Psych, vol 63, p 485-515.

19. Rice, C.E., and Feinstein, S.H., (1943). Sonar System of the Blind: Size Discriminations. Science, vol 148, p 1107-1108.

20. Kay, L. (1980). Air Sonars with Acoustical Display of Spatial Information, in Animal Sonar Systems. R.G. Busnel and J.F. Fish, ed. Plenum Press, New York and London.

21. Riley, L.H. (1966) Evaluation of Sonic Mobility Aid, in Proc. of International Conference on Sensory Devices for the Blind. R. Dufton, ed., St. Dunstans, London, p 153-200.

22. Norris, K.S., Evans, W.E., and Turner, R.N. (1966). Echolocation in an Atlantic Bottlenose Porpoise During Discrimination in Animal Sonar Systems, Biology and Bionics. R.G. Busnel ed. Laboratoire de Physiologie Acoustique INRA-CNRZ, Jouy-en-Josas-78, France.

23. Hammer, C.E., Jr., and Au, W.W.L. (1980). Porpoise Echo Recognition: an Analysis of Controlling Target Characteristics. J Acoust Soc Am, vol 68(5), p 1285-1293.

24. Nachtigall, P.E., Murchison, A.E., and Au, W.W.L. (1980). Cylinder and Cube Shape Discrimination by an Echolocating, Blindfolded Bottlenose Dolphin, in Animal Sonar Systems. R.G. Busnel and J.F. Fish, ed., Plenum Press, New York and London.

25. Au, W.W.L., Schusterman, R.J., and Kersting, D.A. (1980). Sphere-Cylinder Discrimination via Echolocation by Tursiops truncatus, in Animal Sonar Systems. R.G. Busnel and J.F. Fish, ed. Plenum Press, New York and London.

26. Schusterman, R.J., Kersting D.A., and Au, W.W.L. (1980). Response Bias and Attention in Discriminative Echolocation by Tursiops truncatus, in Animal Sonar Systems. R.G. Busnel and J.F. Fish, ed. Plenum Press, New York and London.

27. Au, W.W.L., and Snyder, K.J. (1980). Long Range Target Detection in Open Waters by an Echolocating Atlantic Bottlenose Dolphin. J Acoust Soc Am, vol 68(4) p 1077-1084.

28. Au, W.W.L., and Turl, C.W. (1983). Target Detection in Reverberation by an Echolocating Atlantic Bottlenose Dolphin (Tursiops truncatus). J Acoust Soc Am, vol 73(5), p 1676-1681.

29. Diercks, K.J., Ryan, W.W., Mimeska, E.E., and Weisser, F.L., (1968). Listener Discrimination of Broadband FM Echoes from Simple Geometric Targets. DRL-TM-68-22.

30. Martin, D.W., and Au, W.W.L. (1982). Aural Discrimination of Targets by Human Subjects Using Broadband Sonar Pulses (NOSC Tech. Rep. 847). San Diego: Naval Ocean Systems Center.

31. Shirley, D., and Diercks, K.J. (1970). Analysis of the Frequency Response of Simple Geometric Targets. J Acoust Soc Am, vol 48, p 1275-1282.

32. McLellan, M.E., and Small, A.M. Jr. (1966). Time Separation Pitch Associated with Noise Pulses. J Acoust Soc Am, vol 40(?), p 570-582.

33. Small, A.M. Jr., and McLellan, M.E., (1963). Pitch Associated with Time Delay Between Two Pulse Trains. J Acoust Soc Am, vol 35(8), p 1246-1255.

34. Yost, W.A., and Hill, R. (1978). Strength of Pitches Associated with Ripple Noise. J Acoust Soc Am, vol 64, p 485-492.

35. Ceruti, M.G., Martin, D.W. and Floyd, R.W. (1983). Use of Time-Separation Pitch in Equalizing the Interpulse Intervals of Pulse Triplets by Method of Adjustment. J Acoust Soc Am, vol 73 S-1 S78.

36. Gillespie, R.V. (1964). Time-Separation Pitch: the Effect of Unequal Pulse Amplitudes. University of Iowa Doctoral Dissertation.

37. Chestnut, P.C., and Floyd, R.W. (1981). Aspect-Independent Sonar Target Recognition Method. J Acoust Soc Am, vol 70(3), p 727-734.

38. Au, W.W.L., and Penner, R.H. (1981). Target Detection in Noise by Echolocating Atlantic Bottlenose Dolphins. J Acoust Soc Am, vol 70(3), p 687-693.

39. Scharf, B. (1972). Critical Bands, in Foundations of Modern Auditory Theory, vol I, J.V. Tobias ed., Academic Press, New York.

40. Green, D., Weber, D., and Duncan, J. (1977). Detection and Recognition of Pure Tones in Noise. J Acoust Soc Am, vol 62(4), p 948.

41. The ILS Users Guide (1983). Signal Technology Inc, Goleta, CA.